Claims

We claim:

1. A method for summarizing a compressed video, comprising:

detecting audio peaks in an audio signal of the video;

quantizing motion activity in the video as a continuous stream of pulses; and

correlating the audio peaks with the stream of quantized pulses to identify uninteresting events and interesting events in the video to summarize the video.

2. The method of claim 1 further comprising:

discarding frames of the video associated with the uninteresting events; and

concatenating frames of the video associated with the interesting events to form a summary of the video.

3. The method of claim 1 further comprising:

sub-sampling the audio signal of the video down to a volume contour; and

applying a sliding window to the volume contour to detect a local maximum corresponding to a particular audio peak.

4. The method of claim 3 where the local maximum is detected when (localMax - localMin) > (globalMax - globalMin)/3, using a local minimum, and predetermined global maximum and a predetermined global minimum.

5. The method of claim 3 wherein the sliding window has a duration of one minute, and slides forward in time in half minute steps.

6. The method of claim 1 further comprising:

extracting the motion activity from each P-frame in the video;

applying a moving average filter and a moving median filter to the extracted motion activity to generated smoothed motion activity; and

setting the smoothed motion activity for each P-frame to one if greater than a predetermined threshold, and zero otherwise to quantize the motion activity as the continuous stream of pulses.

7. The method of claim 1 further comprising:

measuring an average of magnitudes of motion vectors of each P-frame to extract the motion activity.

8 The method of claim 6 wherein the predetermined threshold is half a mean motion activity of the compressed video.

9. The method of claim 6 further comprising:

testing each pulse to determine whether the quantized motion activity is at one for at least a first predetermined length of time before falling to zero and remains at zero for a second predetermined length of time; and

selecting the test pulse as a candidate pulse associated with a particular interesting event in the video.

10. The method of claim 9 further comprising:

discarding pulses failing the test from the continuous stream of pulses; and

transforming each candidate pulse to have a third predetermined length of time.

11. The method of claim 10 further comprising:

merging the transformed pulses, time-wise, with the detected audio peaks to obtain a set of time-correlated transformed pulses and audio peaks.

12. The method of claim 11 further comprising:

testing if a rising edge of a particular transformed pulse is less than ten seconds after a particular time-correlated audio peak; and

designating an entire duration starting from the particular audio peak and ending at a first falling edge after the particular audio peak is as a particular interesting event if true.

13. The method of claim 11 further comprising:

testing if a falling edge of a particular transformed pulse is less than two seconds before a particular audio peak; and

designating an entire duration starting from an immediately preceding rising edge and ending at the particular audio peak as a particular interesting event if true.

14. A system for summarizing a compressed video, comprising:

means for detecting audio peaks in an audio signal of the video;

means for quantizing motion activity in the video as a continuous stream of pulses; and

means for correlating the audio peaks with the stream of quantized pulses to identify uninteresting events and interesting events in the video to summarize the video.

15. The system of claim 14 further comprising:

means for discarding frames of the video associated with the uninteresting events; and

means for concatenating frames of the video associated with the interesting events to form a summary of the video.

16. The system of claim 14 further comprising:

means for extracting the motion activity from each P-frame in the video;

means for applying a moving average filter and a moving median filter to the extracted motion activity to generated smoothed motion activity; and

means for setting the smoothed motion activity for each P-frame to one if greater than a predetermined threshold, and zero otherwise to quantize the motion activity as the continuous stream of pulses.

17. The system of claim 16 further comprising:

means for testing each pulse to determine whether the quantized motion activity is at one for at least a first predetermined length of time before falling to zero and remains at zero for a second predetermined length of time; and

means for selecting the test pulse as a candidate pulse associated with a particular interesting event in the video.